BUNDESREPUBLIK DEUTSCHLAND

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 102 29 811.4

Anmeldetag: 3. Juli 2002

Anmelder/Inhaber: Deutsche Telekom AG, Bonn/DE

Bezeichnung: Verschlüsselungsverfahren basierend auf Faktorisierung

IPC: H 04 L 9/30

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 7. Juli 2003
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Faust

Verschlüsselungsverfahren basierend auf Faktorisierung

Zusammenfassung

Bei der Erfindung handelt es sich um ein asymmetrisches Verschlüsselungsverfahren. Der öffentliche Schlüssel besteht aus einer großen zusammengesetzten Zahl n, der geheime Schlüssel besteht aus den Faktoren der zusammengesetzten Zahl. Die Verschlüsselung besteht aus einer Anzahl von Iterationen einzelner Verschlüsselungsschritte, die während der Entschlüsselung sukzessive rückgängig gemacht werden. Die Umkehrung eines einzelnen Verschlüsselungsschrittes erfordert dabei das Lösen einer quadratischen Gleichung modulo m. Der geheime Schlüssel besteht vorzugsweise aus den großen Primzahlen p und q. Der öffentliche Schlüssel ist das Produkt n dieser beiden Primzahlen sowie eine vergleichsweise kleine ganze Zahl L, die größer als eins ist. Die Nachricht m besteht aus zwei ganzzahligen Werten m_1 und m_2, also

$$m = (m_1, m_2)$$

wobei beide Werte in der Menge $Z_n = \{0, 1, 2, \ldots, n-1\}$ liegen.

Die Verschlüsselung geschieht durch die Gleichung

$$c = f^L(m).$$

Verschlüsselungsverfahren basierend auf Faktorisierung

Beschreibung

Die Erfindung betrifft ein asymmetrisches bzw. öffentliches Verschlüsselungsverfahren. Insbesondere betrifft die Erfindung ein Verfahren zur Verschlüsselung von Daten auf der Basis des Faktorisierungsproblems. Hierbei ist die Entzifferung von chiffrierten Daten so komplex wie das Problem, große Primteiler großer Zahlen zu finden. Im Detail sind bei der vorliegenden Erfindung bei der Entschlüsselung quadratische Gleichungen zu lösen.

Um Daten bei der Speicherung oder bei der Übertragung über unsichere Kommunikationskanäle vor dem Zugriff Unbefugter zu schützen, werden Verschlüsselungsverfahren eingesetzt. Dabei werden die Daten so verändert, dass ohne die Kenntnis eines bestimmten Schlüssels diese Veränderung nicht rückgängig gemacht werden kann. Verschlüsselungsverfahren lassen sich in die Kategorien asymmetrische und symmetrische Verschlüsselungsverfahren unterteilen. Bei symmetrischen Verfahren wird derselbe Schlüssel sowohl zur Ver- als auch zur Entschlüsselung verwendet. Asymmetrische Verfahren besitzen zwei unterschiedliche Schlüssel, von denen einer zur Verschlüsselung und der andere zur Entschlüsselung verwendet wird. Dabei kann der Verschlüsselungsschüssel allen Teilnehmern bekannt sein, wohingegen der Entschlüsselungsschlüssel geheim gehalten werden muss. Man bezeichnet daher den Verschlüsselungsschlüssel auch als öffentlichen Schlüssel und den Entschlüsselungsschlüssel als

geheimen Schlüssel. Eine Übersicht über moderne Verschlüsselungsverfahren bietet z. B. das Buch [1] laut Literaturliste.

Bekannt sind die Verfahren von Rabin ([3]) und Williams ([6]), die ebenfalls quadratische Gleichungen verwenden. Allerdings wird bei diesen Verfahren nur die Hälfte der Datenbits pro Übertragung übermittelt. Hierdurch entstehen entsprechende Komplexitätsbeschränkungen und ein höherer Bedarf an Rechenleistung bei der Verschlüsselung und bei der Entschlüsselung.

Das Verfahren von Schwenk und Eisfeld ([5]) bietet bei Polynomen zweiten Grades wenig Sicherheit gegen Angriffe, die Abhängigkeiten der Nachrichtenteile m_1 und m_2 voneinander ausnutzen.

Gelöst wird die Aufgabe durch eine Erfindung mit den Merkmalen der unabhängigen Ansprüche. Hierdurch wird ein asymmetrisches Verschlüsselungsverfahren beschrieben, das auf dem Faktorisierungsproblem basiert. Es hat bei der Verschlüsselung eine geringere Komplexität als das RSA-Verfahren und kann mehr Datenbits pro Verschlüsselung übertragen als das Rabin- bzw. Williamsverfahren.

Wie bereits oben beschrieben wurde, handelt es sich bei der Erfindung um ein asymmetrisches Verschlüsselungsverfahren. Der öffentliche Schlüssel besteht aus einer großen zusammengesetzten Zahl n, der geheime Schüssel besteht aus den Faktoren der zusammengesetzten Zahl. Die Verschlüsselung besteht aus einer Anzahl von Iterationen einzelner Verschlüsselungsschritte, die während der Entschlüsselung sukzessive rückgängig gemacht werden. Die Umkehrung eines einzelnen Verschlüsselungsschrittes erfordert dabei das Lösen

einer quadratischen Gleichung modulo n (siehe unten). Das Lösen einer solchen quadratischen Gleichung ist nur dann leicht möglich, wenn die Faktoren von n bekannt sind.

Der geheime Schlüssel besteht vorzugsweise aus den großen Primzahlen p und q. Der öffentliche Schlüssel ist das Produkt n dieser beiden Primzahlen sowie eine vergleichsweise kleine ganze Zahl L, die größer als eins ist. Die Nachricht m besteht aus zwei ganzzahligen Werten m_1 und m_2, also

$$m = (m_1, m_2)$$

wobei beide Werte in der Menge $Z_n = \{0,1,2, \ldots, n-1\}$ liegen.

Die Verschlüsselung geschieht durch die Gleichung

$$c=f^L(m).$$

Der verschlüsselte Wert c besteht im vorliegenden Fall ebenfalls aus einem Zweitupel ganzer Zahlen aus Z_n, d. h. $c = (c_1, c_2)$.

Die Funktion $f^L(m)$ ist rekursiv definiert durch

$$f^{j+1}(m) = f(f^j(m)),$$

Für $j = 1$ gilt $f^1(m) = f(m) = (f_1(m), f_2(m))$, wobei

$$f_1(m) = m_1+m_2 \bmod n$$

$$f_2(m) = m_1 \cdot m_2 \bmod n.$$

Der verschlüsselte Text wird folglich erhalten mittels der Rekursionen

$$a_{i+1} = a_i+b_i \quad mod\ n \qquad (1)$$

$$b_{i+1} = a_i \cdot b_i \quad mod\ n. \qquad (2)$$

mit den Startwerten $a_0 = m_1$, $b_0 = m_2$ und den Endwerten $c_1 = a_L$, $c_2 = b_L$.

Für die Entschlüsselung muss man die Rekursion umkehren können. Dies geschieht durch Auflösung obiger Gleichungen nach a_i und b_i. Man erhält sogleich die quadratische Gleichung

$$z^2 - a_{i+1} \cdot z + b_{i+1} = 0 \bmod n, \qquad (3)$$

die als Lösungen a_i und b_i besitzt. Auf das Problem der weiteren Lösungen von Gleichung (3) gehen wir später ein. Ist n das Produkt von sehr großen Primzahlen, so ist das Auflösen von quadratischen Gleichungen ohne Kenntnis der Primfaktoren vermutlich ein sehr schwieriges Problem. Bei Kenntnis der Primfaktoren ist dies jedoch leicht möglich. Die gängigen Verfahren zum Wurzelziehen modulo n sind ausführlich in [2] beschrieben.

Damit das Verschlüsselungssystem sicher ist, muss die Rekursion mindestens zweimal durchgeführt werden, da ansonsten bei genau einmaliger Durchführung die Nachrichtenteile m_1 und m_2 linear in den Term $a_1 = m_1 + m_2$ eingehen.

Ein weiterer wichtiger Aspekt ist die Auswahl der korrekten Wurzeln bei der Entschlüsselung

Wenn die Zahl n genau zwei Primfaktoren p und q enthält, hat Gleichung (3) vier Lösungen. Mit wenigen Bits für jedes a_i, i = 1, 2, ... , L kann der Sender dem legitimen Empfänger die Mehrdeutigkeit eliminieren. Zur Auflösung der Mehrdeutigkeit können z. B. von den a_i jeweils Prüf- bzw. Paritätszeichen abgeleitet werden.

Im günstigsten Fall werden, um die Mehrdeutigkeit in jedem Schritt völlig aufzulösen, 2 Bit pro Iterationsschritt

benötigt. Die 4 Lösungen von Gleichung (3) sind gegeben durch

$$z_{i_{1,2,3,4}} = \frac{a_{i+1}}{2} + w_{i_{1,2,3,4}} \bmod n. \qquad (4)$$

wobei

$$w_{i_{1,2,3,4}} = \sqrt{a_{i+1}^2/4 - b_{i+1}} \bmod n$$

die vier Quadratwurzeln des obigen Ausdrucks modulo n sind. Die vier Werte hängen wie folgt zusammen:

$$w_{i_1} = -w_{i_2} \bmod n \quad \text{und} \quad w_{i3} = -w_{i_4} \bmod n$$

Wir wählen die Parität (gerade, ungerade) der vier Wurzeln so, dass

$$w_{i_{1,3}} = \text{gerade und } w_{i_{2,4}} = \text{ungerade}$$

sind.

Eine besonders elegante Lösung, um alle vier Wurzeln voneinander unterscheiden zu können, ist für p≡q≡3 mod 4 wie folgt:

Zusätzlich zur Parität wird als weiteres Unterscheidungskriterium das so genannte Jacobisymbol (w_i/n) benutzt (zur Theorie und zur effizienten Berechnung siehe z. B. [2]). Das Jacobisymbol liefert für nichttriviale Werte von w_i, wie wir sie bei der Dechiffrierung benötigen, den Wert 1 oder −1. Die Berechnung des Jacobisymbols lässt sich mit Aufwand $O(\log^2 n)$ bewerkstelligen.

Die Parität und das Jacobisymbol reichen aus, um genau eine der vier Wurzeln $w_{i_{1,2,3,4}}$ auszuwählen. Die Parität und das Jacobisymbol lassen sich mit 2 Bit codieren. Durch Anhängen

dieser beiden Bits bei jedem der L Iterationsschritte kann man den legitimen Empfänger in die Lage versetzen, die L Iterationsschritte rückgängig zu machen.

Mit w_i wird diejenige Wurzel, die in Gleichung (4) auf die Lösung a_i führt, bezeichnet, also $a_i = a_{i+1}/2 + w_i \bmod n$. Zu dieser Wurzel werden jeweils die Parität und das Jacobisymbol angegeben. Mit dem Wert von a_i folgt dann sofort der Wert für b_i zu $b_i = a_{i+1} - a_i \bmod n$. Zusammenfassend erhält man also

$$a_i = a_{i+1}/2 + w_i \bmod n \qquad (5)$$

$$b_i = a_{i+1}/2 - w_i \bmod n. \qquad (6)$$

Bei der Verschlüsselung wird bei jedem Schritt aus dem Zahlenpaar (a_i, b_i) das Paar (a_{i+1}, b_{i+1}) berechnet sowie die Parität und das Jacobisymbol von $wi = (a_i - a_{i+1}/2) \bmod n$.

Bei Kenntnis der Faktorisierung lassen sich diese Schritte jeweils rückgängig machen durch Auflösung von

$$\sqrt{a_{i+1}^2/4 - b_{i+1} \bmod n},$$

wobei Parität und Jacobisymbol dieser Wurzel dargestellt werden.

Ein weiterer wichtiger Aspekt ist die Parameterwahl. Realistische Größenordnungen für jede der beiden Primzahlen sind derzeit ab ca. 510 Bit, d. h. n hat eine Länge von ca. 1020 Bit. Für L wird eine Größe $O(\log \log n)$ empfohlen, für n von 1000 Bit ein Wert von ca. 3-10.

Die in Zukunft zu wählenden Bitlängen können sich an den Parametern des RSA-Verfahrens orientieren.

Ein Vorteil des hier präsentierten Verfahrens ist, dass die

Anzahl der Nutzdaten doppelt so groß wie bei vergleichbaren Verfahren ist.

Mit Standardalgorithmen wird eine Verschlüsselungskomplexität von $O(L \log^2 n)$ erreicht, wenn man den Aufwand für eine Multiplikation mit $O(\log^2 n)$ rechnet. Für die Entschlüsselungskomplexität muss man bei Benutzung von gängigen Algorithmen mit einem Aufwand von $O(L \log^3 n)$ rechnen. Wählt man für L eine Größenordnung von $O(\log \log n)$, so ergibt sich bei der Verschlüsselung gegenüber dem RSA-Verfahren ein Zeitvorteil (neben der größeren Nutzdatenrate).

Wie beim Rabin- und Williamsverfahren muss man bei der Realisierung beachten, dass jeweils nur die richtigen Wurzeln von Gleichung (3) bei der Entschlüsselung den Dechiffrierer verlassen, da ansonsten die Zahl n faktorisiert werden kann.

In einer weiteren Ausgestaltung wie beim RSA-Verfahren kann der Modul n auch mehr als zwei große Primfaktoren enthalten. Dementsprechend erhöht sich natürlich auch die Anzahl der Lösungen von Gleichung (3).

Eine weitere Verallgemeinerung wird dadurch erreicht, dass bei der Rekursion noch zusätzliche Konstanten eingeführt werden:

$$a_{i+i} = k_1 \cdot a_i + k_2 \cdot b_i \bmod n$$

$$b_{i+1} = k_3 \cdot a_i \cdot b_i \bmod n,$$

die als Teil des öffentlichen Schlüssels bekannt gemacht werden. Die Dechiffrierung geschieht in entsprechend modifizierter Form.

In einer weiteren Ausführungsform wird die Größe der Tupel verändert. Statt mit Zweitupeln $m = (m_1, m_2)$ kann man auch mit q Tupeln arbeiten. Im Folgenden wird die Erweiterung anhand

von Drei-Tupeln illustriert. Die Nachricht besteht nun aus dem Dreitupel

$$m = (m_1, m_2, m_3)$$

Die Formel für den *L*-ten Iterationsschritt lautet unverändert

$$f^{j+1}(m) = f(f^j(m)),$$

wobei allerdings die Grunditeration $f^1(m) = (f_1(m), f_2(m), f_3(m))$ wie folgt gebildet wird

$$f_1(m) = m_1 + m_2 + m_3 \bmod n$$

$$f_2(m) = m_1 \cdot m_2 + m_1 \cdot m_3 + m_2 \cdot m_3 \bmod n$$

$$f_3(m) = m_1 \cdot m_2 \cdot m_3 \bmod n \,.$$

Die Rückrechnung erfolgt durch Auflösung einer Gleichung dritten Grades. Die Unterscheidung der Wurzeln kann wieder durch entsprechend von den Zwischenergebnissen abgeleiteten Informationen (Paritäts-, Jacobisymbol, etc.) geschehen. Die Erweiterung auf Grade größer oder gleich vier kann in analoger Weise geschehen. Bei der Iteration sind im Wesentlichen die elementarsymmetrischen Newtonschen Terme zu betrachten, zu denen noch zusätzliche Konstanten, wie bereits oben beschrieben wurde, hinzutreten können.

Im Folgenden wird anhand eines Beispiels das Verfahren der vorliegenden Erfindung deutlich gemacht. Die im Folgenden gewählten Zahlen sind aus Gründen der Übersichtlichkeit sehr klein gewählt. Sei $n = 8549 = p \cdot q$, mit den geheimen Primzahlen $p = 83$ und $q = 103$. Die Anzahl der Iterationen sei $L = 3$ und die zu verschlüsselnde Nachricht sei gegeben *durch* $m = (m_1, m_2) = (123,456)$. Gerade Parität werde durch eine Null,

ungerade Parität durch eine Eins codiert. Hierzu dient das Paritätsbit b_P. Ist das Jacobisymbol gleich eins wird eine Eins, ist es gleich minus eins, wird eine Null codiert. Hierzu wird das Jacobibit b_J benutzt.

Man erhält die folgenden Werte

$$(a_0,\ b_0) = (123,456)$$

$$(a_1,\ b_1) = (579,4794)$$

$$(a_2,\ b_2) = (5373,5850)$$

$$(a_3,\ b_3) = (2674,5926)$$

Zu jedem der drei Paare $(a_1,\ b_1)$, $(a_2,\ b_2)$ und $(a_3,\ b_3)$ werden noch $L \cdot 2$ Bits Paritäts- und Jacobibits, die im Beispiel durch den folgenden Binärvektor $(b_{P_3}, b_{J_3}, b_{P_2}, b_{J_2}, b_{P_1}, b_{J_1}) = (0,0,1,1,0,1)$ gegeben sind, hinzugefügt.

Der Empfänger bestimmt zunächst die vier Wurzeln $w_{2_{1,2,3,4}}$ = 1629,4036,4513,6920. Anhand von $b_{P_3} = 0$ erkennt er, dass die richtige Wurzel gerade ist. Es bleiben also nur 4036 und 6920. Von diesen ist $(4036/8549) = -1$ und $(6920/8549) = 1$. $b_{J_3} = 0$ besagt, dass 4036 die richtige Wahl ist. Analoges Vorgehen führt zur vollständigen Entschlüsselung.

Auf die Mitübertragung der Bits zur Auflösung der Mehrdeutigkeit kann man in bestimmten Anwendungsfällen verzichten, z. B. wenn die unverschlüsselte Nachricht m Redundanz enthält. Dies ist beispielsweise bei normalen Texten der Fall oder wenn bereits in m ein so genannter Hashwert untergebracht wurde. Dies wird jedoch durch einen um den Faktor 4^L erhöhten Entschlüsselungsaufwand erkauft. Entsprechende Kompromisse sind ebenfalls möglich, z. B.

verringert die Angabe von nur der Parität in jedem der L Schritte die Zahl der mitzusendenden Bits auf L Bit und erhöht den Entschlüsselungsaufwand um den Faktor 2^L.

Wie bei den in der Literatur ([1],[3],[4],[5]) bekannten asymmetrischen Verfahren kann man auch bei dem vorgeschlagenen Verfahren im Wesentlichen durch Vertauschen von Ver- und Entschlüsselungsoperationen ein so genanntes digitales Signaturverfahren erhalten.

Liste der zitierten Literatur:

[1] A. J. Menezes, P. C. van Oorschot, S. A. Vanstone, "Handbook of Applied Cryptography", CRC Press, 1996.

[2] E. Bach, J. Shallit, "Algorithmic Number Theory", Vol. 1, Efficient Algorithms, The MIT Press, Cambrigde, Massachusetts, 1996.

[3] M. O. Rabin, "Digitalized Signatures and Public-Key Functions as as intractable as Factorization ", MIT/LCS/TR-212, 1979.

[4] R. L. Rivest, A. Shamir, L. Adleman, "A Method for Obtaining Digital Signatures and Public Key Cryptosystems", Communications of the ACM, Vol. 21 Nr.2, pp. 120-126, Feb. 1978.

[5] J. Schwenk, J. Eisfeld, "Public Key Encryption and Signature Schemes based and Polynomials over Z_n", Eurocrypt 1996, LNCS 1070, Springer-Verlag Berlin Heidelberg 1996. V

[6] H. Williams, "A Modification of the RSA Public-Key Equation Procedure ", IEEE Transactions on Information Theorie, Vol. IT-26, No. 6, November 1980. V

Patentansprüche

1. Verfahren zur Verschlüsselung von Daten nach einem asymmetrischen Verfahren, basierend auf einem Faktorisierungsproblem, mit einem öffentlichen Schlüssel und einem privaten Schlüssel, wobei der Schlüssel die Iterationszahl L sowie die öffentliche Schlüssel die Iterationszahl L sowie die zusammengesetzte Zahl n ist, wobei n vorzugsweise das Produkt mehrerer großer Primzahlen ist, wobei der private Schlüssel aus der Faktorisierung von n besteht, wobei die zu verschlüsselnde Nachricht $m = (m_1, m_2)$ mindestens aus den Bestandteilen m_1 und m_2 besteht, wobei eine Verschlüsselungsfunktion $f(x)$ insgesamt L mal iteriert wird, mit $c=(c_1,c_2)=f^L(m)$, wobei $f(m)=(f_1(m),f_2(m))$ gilt und $f_1=(m_1 op_1 m_2) \bmod n$ sowie $f_2=(m_1 op_2 m_2) \bmod n$, wobei op_1 vorzugsweise eine Addition ist und op_2 vorzugsweise eine Multiplikation ist, wobei die Verschlüsselungsfunktion $f(x)$ so gewählt ist, dass durch L-malige Auflösung einer quadratischen Gleichung modulo n die Verschlüsselungsiteration rückgängig zu machen ist, wodurch aus der verschlüsselten Information $c = (c1, c2)$ die ursprüngliche Nachricht wiederzugewinnen ist.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass eine Mehrdeutigkeit der quadratischen Gleichung durch zusätzliche Bits von a_i und b_i, eliminiert wird.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, dass die Mehrdeutigkeit der quadratischen Gleichung durch Berechnung einer Parität und eines Jacobisymbols eliminiert werden, die insbesondere bei Primzahlen der Form 3 mod 4 durch 2 Bit je Iterationsschritt mitgeteilt werden können.

4. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass allgemeine Iterationen $f_1=(k_1 \bullet m_1 + k_2 \bullet m_2) \bmod n$ sowie $f_2 = k_3 \bullet m_1 \bullet m_2 \bmod n$ verwendet werden, wobei die Konstanten Teil des öffentlichen Schlüssels sind.

5. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die zusammengesetzte Zahl n als öffentlicher Schlüssel mehr als zwei Faktoren enthält.

6. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche , dadurch gekennzeichnet, dass die Nachricht nun aus einem N-Tupel besteht $m=(m_1 \dots m_N)$, wobei die Formel für den *L*-ten Iterationsschritt in jedem Iterationsschritt Abhängigkeiten von N Werten verwendet.

7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, dass die Mehrdeutigkeit durch zusätzliche Bits aufgelöst wird, die aus den in jeder Iteration erhaltenen Werten abgeleitet werden.

8. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Mehrdeutigkeit durch Redundanz in den übermittelten Daten aufgelöst wird.

9. Verfahren zur Erzeugung einer Signatur, dadurch gekennzeichnet, dass durch Vertauschung der Ver- und Entschlüsselungsschritte aus dem vorhergehenden Verfahren eine Signatur erzeugt wird.

10. Software für einen Computer, dadurch gekennzeichnet, dass ein Verfahren nach einem oder mehreren der vorhergehenden Ansprüche implementiert ist.

11. Datenträger für einen Computer, gekennzeichnet durch die Speicherung einer Software nach dem vorhergehenden Softwareanspruch.

12. Computersystem, gekennzeichnet durch eine Einrichtung, die den Ablauf eines Verfahrens nach einem oder mehreren der vorhergehenden Verfahrensansprüche erlaubt.